

## 1 METHOD OF DNA SEQUENCING USING CLEAVABLE TAGS

2 Background of the Invention

3 DNA sequencing is an important analytical technique critical to generating  
4 genetic information from biological organisms. The increasing availability of rapid  
5 and accurate DNA sequencing methods has made possible the determination of the  
6 DNA sequences of entire genomes, including the human genome. DNA sequencing  
7 has revolutionized the field of molecular biological research. In addition, DNA  
8 sequencing has become an important diagnostic tool in the clinic, where the rapid  
9 detection of a single DNA base change or a few base changes can be used to detect,  
10 for example, a genetic disease or cancer.

11 Most current methods of DNA sequencing are based on the method of Sanger  
12 (*Proc. Natl. Acad. Sci. U.S.A.*, 74, 5463 (1977)). This method relies on gel  
13 electrophoresis of single stranded nucleic acid fragments that are generated when a  
14 polymerization-extension reaction of a primer is terminated by incorporation of a  
15 radioactively labeled dideoxynucleotide-triphosphate. Short strands of DNA are  
16 synthesized under conditions that produce DNA fragments of variable length using a  
17 DNA polymerase and deoxynucleotide triphosphates (dNTP). A small amount of  
18 dideoxynucleotide triphosphates (ddNTP) is introduced into the DNA synthesis  
19 mixture so that chain terminating ddNTPs are sometimes integrated into a growing  
20 strand. Typically, four different extension reactions are performed side by side, each  
21 including a small amount of one ddNTP. Each extension reaction produces a mixture  
22 of DNA fragments of different lengths terminated by a known ddNTP. The ratio of  
23 ddNTPs to dNTPs is chosen so that the populations of DNA fragments in any given

1 extension reaction includes fragments of all possible lengths (up to some maximum)  
2 terminating with the relevant ddNTP. The nucleic acid fragments are separated by  
3 length in the gel, typically utilizing a different lane in a polyacrylamide gel for each of  
4 the four terminating nucleotide bases being detected. However, such size exclusion  
5 chromatography is generally a low resolution method limited to reading short  
6 sequences.

7 A variation of this method utilizes dyes rather than radioactivity to label the  
8 ddNTPs. Different dyes are used to uniquely label each of the different ddNTPs (i.e.,  
9 a different dye may be associated with each of A, G, C, and T termination) (Smith et  
10 al. and Prober et al. *Science* 238:336-341, 1987). In the method of Smith, fluorescent  
11 dyes are attached to the 3' end of the dNTP converting it into a ddNTP. The use of  
12 four different dye labels allows the entire sequencing reaction to be conducted in a  
13 single reaction vessel and results in a more uniform signal response for the different  
14 DNA fragments. The dye-terminated dNTPs are also able to be electrophoresed in a  
15 single lane. The advent of capillary electrophoresis further increased the separation  
16 efficiency of this method, allowing shorter run times, longer reads, and higher  
17 sensitivity.

18 Despite these advances, DNA sequencing methods that rely on electrophoresis  
19 to resolve DNA fragments according to their size are limited by the rate of the  
20 electrophoresis and the number of bases that are detectable on the gel. In addition,  
21 real time imaging of the gel is not possible. Accordingly, in order to increase the  
22 speed and reliability of the sequencing reaction, great effort has been made to  
23 automate these steps. Automated DNA sequencing machines are now available that  
24 are capable of high throughput sequencing for both genomic sequencing and routine

1 clinical applications. However, these newer techniques remain cumbersome,  
2 requiring specialized chemicals and the intensive labor of skilled technicians.

3 One newer method of DNA sequencing, "pyrosequencing" or "sequencing-by-  
4 synthesis," disclosed in WO 98/13523, is based on the concept of detecting inorganic  
5 pyrophosphate (PPi), which is released during a polymerase reaction. As in the  
6 Sanger method, a sequencing primer is hybridized to a single stranded DNA template  
7 and incubated with a DNA polymerase. In addition to the polymerase, the enzymes  
8 ATP sulfurylase, luciferase, and apyrase, and the substrates, adenine 5'  
9 phosphosulfate (APS) and luciferin, are added to the reaction. Subsequently,  
10 individual nucleotides are added. When the added nucleotide is complementary to the  
11 next available base in the template strand, it is incorporated into the extension  
12 product. Such incorporation of a complementary base is accompanied by release of  
13 pyrophosphate (PPi), which is converted to ATP in the presence of adenosine 5'  
14 phosphosulfate by apyrase in a quantity equimolar to the amount of incorporated  
15 nucleotide. The ATP generated by the reaction with apyrase then drives the luciferase  
16 mediated conversion of luciferin to oxyluciferin, generating visible light in amounts  
17 that are proportional to the amount of ATP and thus the number of nucleotides  
18 incorporated into the growing DNA template. The light produced by the luciferase-  
19 catalyzed reaction is detected by a charge coupled device (CCD) camera and detected  
20 as a peak in a pyrogram<sup>TM</sup>.

21 In a pyrosequencing reaction, if the first nucleotide added to the reaction is not  
22 complementary to the next available nucleotide on the growing DNA strand there is  
23 no light generated. If no light is generated by the addition of the first nucleotide, a  
24 second of four dNTPs is added sequentially to the reaction to test whether it is the

1 complementary nucleotide. This process is continued until a complementary  
2 nucleotide is added and detected by a positive light read-out. Whether or not a  
3 positive light reaction is generated, apyrase, a nucleotide-degrading enzyme,  
4 continuously degrades unincorporated dNTPs and excess ATP in the reaction mixture.  
5 When degradation is complete, another dNTP is added.

6 Although pyrosequencing is capable of generating high quality data in a  
7 relatively simple fashion, this method has several drawbacks. First, the productivity  
8 of the method is not high, reading only about 1 base per 100 seconds. The rate of the  
9 reaction is limited by the necessity of having to add new enzymes with each addition  
10 of the dNTPs in addition to the necessity of having to test each of the four dNTPs  
11 separately. In addition, it has been found that the dATP used in the chain extension  
12 reaction interferes in subsequent luciferase-based detection reactions by acting as a  
13 substrate for the luciferase enzyme. Finally, these reactions are expensive to run.

14 While pyrosequencing improves the ease and speed with which DNA  
15 sequencing is achieved, there exists the need for improved sequencing methods that  
16 allow more rapid detection. Preferred techniques would be amenable to automation  
17 and allow the sequence information to be revealed simultaneously with or shortly  
18 after the chain extension reaction.

#### 19 **Summary of the Invention**

20 The present invention provides a novel system for sequencing nucleic acid  
21 molecules. In particular, the invention utilizes dNTPs that are 3' end labeled with a  
22 cleavable tag that distinguishes the dNTP from other dNTPs (e.g., the tag may be  
23 unique to the dNTP). The cleavable tags are functional groups that can be later

1 removed by any appropriate means, including but not limited to, exposure to chemical  
2 cleavage conditions or light. dNTPs labeled with the cleavable tags function as  
3 terminated dNTPs (cdNTPs), in that their incorporation into a single stranded nucleic  
4 acid molecule via a primer extension reaction blocks further extension. However,  
5 removal of the tag converts the cdNTP back into an extendible nucleotide.

6 According to the present methods, a sequencing primer is hybridized to a  
7 nucleic acid template, e.g., a single stranded DNA template, and incubated with an  
8 enzyme (DNA polymerase) and four cdNTPs (tag terminated dATP (cdATP), dCTP  
9 (cdCTP), dGTP (cdGTP), and dTTP (cdTTP)). The DNA polymerase then extends  
10 the primer by adding to it whichever cdNTP is complementary to the next available  
11 base on the template strand. Only a single cdNTP is incorporated, because the cdNTP  
12 cannot be further extended.

13 After completion of a single base addition, unreacted (excess) cdNTPs are  
14 removed from the reaction mixture, which includes the extended primer, the DNA  
15 polymerase, and the single stranded DNA template. The step of removing can be  
16 accomplished by any of a variety of means that would be apparent to one skilled in  
17 the art. For example, if the reaction mixture is contained in a chamber that has an  
18 attached membrane (e.g., an ultrafiltration membrane that allows small molecules  
19 such as water, salts, and cdNTPs to pass through, but does not allow passage of large  
20 molecules such as single stranded DNA), the excess cdNTP can be washed through  
21 the membrane. Alternatively, if the single stranded DNA is attached to a solid  
22 support, the excess cdNTPs can be washed away from the single stranded DNA  
23 without dislodging the hybridized, extended primer.

1           Once the step of removing is complete, the tag is cleaved from the cdNTP that  
2 is extended into the single stranded DNA template. In certain embodiments, the  
3 cleavage occurs by photo-cleavage of the tag from the extended single stranded DNA  
4 template by exposure to light. Alternatively, in other preferred embodiments, the  
5 cleavage occurs by exposure of the single stranded DNA template to a chemical  
6 cleaving agent, e.g., an acid or a base. Whichever cleavage method is employed, the  
7 result is liberation of the 3' end of the extension product for further extension.

8           The cleaved tag is then washed through the membrane into a detector for  
9 identification, thereby identifying the complementary base in the single stranded  
10 DNA template and determining the DNA sequence. The detector used to identify the  
11 tag is chosen based on the type of cleavable tag employed. Any of a variety of tags  
12 may be employed in the present invention, as would be recognized by the skilled  
13 artisan, and such tags are described herein. Once the tag is cleaved, the four cdNTPs  
14 are added back to the primer extension reaction mixture and the cycle of extension,  
15 tag cleavage, and identification is repeated.

16           In other preferred embodiments, short oligonucleotides are employed in a  
17 ligation reaction to determine the sequence of a particular DNA sample. The  
18 sequence of a DNA sample is determined by incorporating "X" complementary bases  
19 (e.g., 2mers, 3mers, or more) at a time onto the single stranded DNA template  
20 adjacent to a primer using a DNA ligase instead of using a DNA polymerase. Each  
21 oligonucleotide is tagged and labeled with a cleavable tag so that the position of each  
22 base in the sequence of the oligonucleotide can be identified. The tag further prevents  
23 ligation of the oligonucleotides to one another.

1 According to this aspect of the invention, a template DNA is exposed to the  
2 oligonucleotides, the oligonucleotides are allowed to hybridize to the template DNA,  
3 and a ligation reaction is allowed to take place on the DNA template such that one  
4 complementary oligonucleotide is incorporated onto the DNA template adjacent to the  
5 annealed primer. Following ligation, the unincorporated oligonucleotides are washed  
6 away from the DNA sample and the tags are cleaved and analyzed to determine the  
7 nucleic acid sequence.  
8

#### 9 Detailed Description of Certain Preferred Embodiments

10 The present invention provides a system for sequencing a DNA molecule  
11 using deoxynucleotide triphosphates of adenine, thymine, guanine, and cytosine that  
12 are each labeled with a different cleavable tag that is used to identify the base. In  
13 preferred embodiments, the cleavable tag further acts as a terminator to extension of a  
14 single stranded DNA template in a polymerase extension reaction until the tag is  
15 removed from the incorporated base. Once removed, the tag is isolated and identified,  
16 and the process of base addition and cleavage is repeated. More particularly, the steps  
17 of extension, cleavage, and detection are repeated until sufficient sequence of the  
18 single stranded DNA template is determined.

19 According to certain preferred embodiments, inventive methods of  
20 determining the sequence of a nucleic acid include the steps of (a) hybridizing an  
21 oligonucleotide to a single stranded DNA, wherein the oligonucleotide is  
22 complementary to at least a portion of the single stranded DNA; (b) providing a DNA  
23 polymerase and four deoxynucleotide triphosphates (dNTPs) (e.g., dATP, dGTP,

1 dCTP, and dTTP) wherein each dNTP is 3' end labeled with a cleavable tag (cdNTP)  
2 that distinguishes the dNTP from other dNTPs; (c) extending the single stranded DNA  
3 hybridized to the oligonucleotide by adding one complementary cdNTP in a  
4 polymerase extension reaction, wherein the tag on the extended cdNTP blocks further  
5 extension by the DNA polymerase; (d) optionally removing excess cdNTPs that are  
6 not extended onto the single stranded DNA; (e) cleaving the tag from the extended  
7 cdNTP; and (f) detecting the tag so that the incorporated base is detected. In certain  
8 preferred embodiments, the method includes the step of repeating steps (a) through (f)  
9 on the sample of single stranded DNA.

10 As indicated above, prior to cleavage of the tag from the extended base on the  
11 DNA template, the excess, unincorporated cdNTPs are preferably removed from the  
12 extension reaction. According to the invention, the tags may be removed by any of a  
13 variety of washing or rinsing procedures that separate the excess, unincorporated  
14 dNTPs from the extended DNA template. In one preferred embodiment, the  
15 extension reaction is contained within a chamber that has an attached filtration  
16 membrane, e.g., an ultrafiltration membrane, that allows small molecules such as  
17 water, salts, and cdNTPs to pass through, while retaining large molecules such as  
18 ssDNA. According to this particular embodiment, a wash solution, e.g., a buffered  
19 saline solution such as phosphate buffered saline, is passed through the ultrafiltration  
20 membrane of the chamber containing the oligonucleotide primer, the DNA  
21 polymerase, the cdNTPs, and the extended DNA to rinse away the excess cdNTPs.  
22 Alternatively, if the DNA template is attached to a solid support, a wash solution may  
23 be passed over the solid support to rinse the excess cdNTPs away from the solid  
24 support.

1           In a related embodiment, the sequencing method of the present invention is  
2   also amenable to sequence determination via oligonucleotide ligation. This technique  
3   requires first exposing the DNA template to a collection of tagged oligonucleotides  
4   (e.g., the tagged oligonucleotides may be a collection of short randomized  
5   oligonucleotides). Preferably, a 3' tag blocks further ligation at the 3' end of the  
6   oligonucleotide to other oligonucleotides in the collection. However, it will be  
7   appreciated that if the tag is located at a position on the oligonucleotide other than the  
8   3' end, the 3' end of the oligonucleotide would still need to be blocked, for example,  
9   with another functional group. Once the DNA template is mixed with the tagged  
10   oligonucleotides, the oligonucleotides are allowed to hybridize to the DNA template  
11   in a position adjacent to an oligonucleotide primer so that the oligonucleotide and  
12   primer can be ligated. Unligated oligonucleotides are then rinsed away from the DNA  
13   template, tags are cleaved from the ligated oligonucleotides, and cleaved tags  
14   representing the bases of the ligated oligonucleotide are detected. This cycle can be  
15   repeated as described, with addition of the oligonucleotide mix occurring at each  
16   repetition.

17           In certain preferred embodiments, the number of tags attached to the 3' end of  
18   the oligonucleotide may be based on the sequence length of the oligonucleotide. For  
19   example, an oligonucleotide that is three bases long may be 3' end labeled with three  
20   tags that are attached in a sequential order matching the sequential order of the bases  
21   of the oligonucleotide.

22           In the oligonucleotide ligation reaction, as with the polymerase reaction, the  
23   DNA template is a single stranded DNA template that is annealed to a primer for  
24   primer extension. By "single stranded DNA template" is meant any single stranded

1 DNA template or single stranded DNA template that is partially single stranded, i.e.,  
2 may be partially double stranded. In one preferred embodiment, an oligonucleotide  
3 that is 3' end blocked and complementary to the sequence adjacent to the primer  
4 anneals to the DNA template and is joined to the adjacent primer via a ligase (e.g., T4  
5 DNA ligase). The tags on the complementary oligonucleotide are then removed for  
6 detection and identification, freeing the 3' end of the complementary oligonucleotide  
7 for subsequent rounds of ligation. In such subsequent rounds, the ligase joins the next  
8 complementary blocked oligonucleotide to the 3' end of the previously extended  
9 primer and the cycle repeats.

10 As mentioned above, the collection of oligonucleotides may include short  
11 randomized oligonucleotides. Those skilled in the art will appreciate that the longer  
12 the oligonucleotide, the greater the number of oligonucleotides will have to be  
13 generated to encompass all possible random oligonucleotide sequences, based on  
14 randomization between four bases at each position of the oligonucleotide. For  
15 example, generation of a collection of 2mers that encompasses all possible 2mers  
16 would require sixteen oligonucleotide sequences; generation of a collection of 3mers  
17 that encompasses all possible 3mers would require a panel of 64 oligonucleotide  
18 sequences; 4mers would require a panel of 256 oligonucleotide sequences, etc.  
19 Identification of an optimal oligonucleotide length may require simply testing various  
20 short random oligonucleotide mixes and determining which give the most rapid and  
21 accurate DNA sequencing results via oligonucleotide ligation. Of course, the longer  
22 the oligonucleotide, the faster the sequencing reaction will proceed, due to the  
23 increased number of incorporated bases detected simultaneously. Using this

1 approach, at each round of the sequencing reaction, the oligonucleotide sequence that  
2 is ligated to the primer is detected and identified.

3 In certain preferred embodiments, it is conceivable that only a small subset of  
4 all possible oligonucleotides need to be used in the sequencing reaction, for example,  
5 if the sequence of the DNA template were partially determined (i.e., if certain  
6 positions of the oligonucleotide were fixed, fewer base positions would need to be  
7 randomized, limiting the number of oligonucleotide required to include all possible  
8 permutations). In this particular embodiment, the oligonucleotides could be longer  
9 (e.g., 5mer, 6mers, 7mers, 8mers, 9mers, 10mers, or greater than 15mers, 20mers,  
10 25mers, 30mers and higher). It is also possible that in certain circumstances one  
11 would not need to use as many tags i.e., one would not need to use one tag for every  
12 base. For example, one unique tag could be used to identify an entire oligonucleotide  
13 sequence.

14 In other preferred embodiments, two or more unique tags could be used to  
15 identify an entire oligonucleotide sequence, the total number of tags being less than  
16 the total number of bases in the oligonucleotides (e.g., each tag could identify short  
17 sequential stretches of oligonucleotides (e.g., a 3mer or a 4mer etc.) within the entire  
18 oligonucleotide sequence). In a related embodiment, an oligonucleotide, particularly  
19 an oligonucleotide used in the ligation aspect of the invention, may not be randomized  
20 at every position (e.g., if certain nucleotide positions are fixed), and may even be  
21 randomized at only one or several positions, e.g., 1-2, 1-3, 1-4 or 1-5 positions.  
22 Under these circumstances, only a subset of possible variations would be relevant.

23 In embodiments where the length of the oligonucleotide sequence increases  
24 the number of tags required to identify the oligonucleotide sequence, the availability

1 of many unique mass tags makes mass spectrometry a particularly useful system for  
2 detection. Since each short random oligonucleotide must be labeled with a unique  
3 tags, the short random oligonucleotide may have a maximum length in certain  
4 circumstance (e.g., the length and number of oligonucleotides in a collection of  
5 oligonucleotides may be limited by the availability of different unique tags).  
6 However, mass tags may have the same nominal mass and vary in structure, thereby  
7 increasing the diversity of tags available.

8 Although the level of diversity available in the mass spectrometry system is  
9 sufficient to permit unique MS/MS fragmentation, those skilled in the art will  
10 appreciate that, because identification of the incorporated oligonucleotides is based on  
11 the MS/MS parent/daughter transition, if an MS/MS approach is used, multiplexing  
12 target DNA samples is not possible. The MS/MS approach requires the isolation of a  
13 single mass followed by fragmentation and mass analysis. Multiplexing would  
14 present too many masses for isolation and fragmentation to be practical. However,  
15 the MS/MS approach would be helpful in increasing the potential number of mass  
16 tags required for coding the oligonucleotides used in the ligation reaction.

17 Thus, according to the oligonucleotide ligation aspect of the invention, the  
18 sequence of a single stranded DNA template may be determined by (a) hybridizing a  
19 complementary oligonucleotide to a single stranded DNA adjacent to a primer,  
20 wherein the oligonucleotide is 3' end labeled with one or more cleavable tags unique  
21 to the oligonucleotide sequence; (b) ligating the hybridized complementary  
22 oligonucleotide to the primer, wherein the one or more tags on the extended cdNTP  
23 blocks further ligation by the DNA ligase; (c) optionally removing excess  
24 oligonucleotides that are not ligated; (d) cleaving the one or more tags from the

1 ligated complementary oligonucleotide; (e) detecting the one or more tags. In certain  
2 preferred embodiments, steps (a) through (e) are repeated on the single stranded  
3 DNA.

4       As in the polymerase reaction, in the ligation reaction, prior to cleavage of the  
5 tag(s) from the extended oligonucleotide on the DNA template, the excess  
6 oligonucleotides are preferably removed. The oligonucleotides may be removed by  
7 any of a variety of separation procedures that may include washing or rinsing the  
8 unincorporated oligonucleotides away from the extended DNA template. As with the  
9 polymerase reaction, in one preferred embodiment, the ligation reaction is contained  
10 within a chamber that has an attached filtration membrane that would allow short  
11 oligonucleotides to pass through, while retaining larger molecules, such as the DNA  
12 template. Alternatively, the DNA template is attached to a solid support and a wash  
13 solution may be passed over the solid support to remove the unincorporated tagged  
14 oligonucleotides.

15       As will be appreciated by those skilled in the art, whether the sequencing  
16 reaction employs a DNA polymerase or a DNA ligase, any tag that is cleavable by  
17 chemical means or by light can be used in the present invention. In certain preferred  
18 embodiments, the tag is cleaved by exposure to an acid or a base. In other preferred  
19 embodiments, the tag is cleaved by exposure to light, i.e., in a photo-cleavage  
20 reaction. The cleavable tags themselves include any functional group that imparts a  
21 unique identity onto the oligonucleotide or base that is tagged. According to the  
22 present invention, useful tags include, e.g., fluorescent tags, mass tags, IR tags, UV  
23 tags, potentiometric tags, etc. For example, a fluorescent tag may be attached to a  
24 dNTP prior to the primer extension reaction, and then may be cleaved from the dNTP

1 after the dNTP is incorporated into the extended DNA strand by exposure of the  
2 extended DNA strand to an acid, a base, or light, and analyzed using fluorescence  
3 spectrometry. As but another example, a base having an acid, base, or light cleavable  
4 mass tag, after incorporation into the DNA template, may be cleaved from the  
5 extended DNA strand using the appropriate cleavable agent, and then may be  
6 analyzed using mass spectrometry.

7 The DNA sequencing methods of the present invention provide an advantage  
8 over existing Sanger-based methods by eliminating the need to separate cDNA  
9 fragments on a gel, resulting in longer sequence reads. The present method is rapid  
10 and fully automatable. In addition, the selection and detection of one of the four  
11 bases is carried out simultaneously.

12 Alternatively, the identification step need not be carried out simultaneously  
13 with the cycling of the reaction. For example, the tags from each cycle may be  
14 collected and pooled (e.g., onto a 96 well plate). Alternatively, the tags from each  
15 cycle may be spatially arrayed (e.g., onto a chip) and the positional information used  
16 for identification. Using either method, the tags are analyzed subsequent to the  
17 cycling reaction by art available means. Such collection and analysis may increase  
18 the speed of the sequencing reaction to increase the throughput of the technique. Of  
19 course one skilled in the art would recognize that the appropriate instrumentation is  
20 required to analyze the collected tags.

21 Certain aspects of the present invention are described in further detail below.

22

23

24

1    *Nucleic Acid Preparation*

2            In certain preferred embodiments of the invention, the DNA sample is a single  
3    stranded DNA template. Alternatively, if in a polymerase extension reaction a  
4    thermostable DNA polymerase enzyme is employed, the DNA sample may be double  
5    stranded.

6            The DNA sample of the invention may be provided from any available source  
7    of DNA, including, for example, a biological sample, including not only samples  
8    obtained from living organisms (e.g., mammals, fish, bacteria, parasites, viruses,  
9    fungi, and the like) or from the environment (e.g., air water, or solid samples), but  
10   biological materials which may be artificially or synthetically produced (e.g., phage  
11   libraries, organic molecule libraries, pools of genomic clones, and the like).  
12   Representative examples of biological samples include biological fluids (e.g., blood,  
13   semen, cerebral spinal fluid, urine), biological cells (e.g., stem cells, B or T cells,  
14   fibroblasts, and the like), and biological tissues. Alternatively, the DNA may be a  
15   cDNA synthesized from an RNA sample (e.g., from a natural or synthetic source).  
16   Such cDNA synthesis may be carried out using reverse transcription, and such  
17   systems are readily available.

18           The DNA sample, whether from a biological or synthetic source, may further  
19   be amplified, particularly if the amount of sample DNA is small. Amplification can  
20   be carried out by any art available method, for example, *in vitro* by PCR or Self  
21   Sustained Sequence Replication (3SR) or *in vivo* using a vector. Alternatively, if  
22   desired, *in vitro* and *in vivo* amplification may be used in combination (see, e.g.,  
23   McPherson, "PCR: A Practical Approach," Oxford University Press, New York,

1 1991). Within other embodiments of the invention, the DNA samples of the present  
2 invention may be generated by, for example, a ligation or cleavage reaction.

3 According to the invention, the DNA sample, amplified or unamplified, is  
4 either immobilized on a solid support or in solution. In the case of an amplified DNA  
5 sample, those skilled in the art will recognize that any amplification procedure may be  
6 modified to allow for attachment of the amplified DNA sample to a solid support. For  
7 example, a chosen PCR primer may be immobilized to a solid support or may be  
8 provided with a means for attachment to a solid support. Immobilization may take  
9 place as part of a PCR amplification, e.g., where one or more primers is attached to a  
10 support. Alternatively, one or more primers may carry a functional group, e.g., a  
11 biotin or thiol group, permitting subsequent immobilization of the DNA sample.  
12 Immobilization of the 5' end of a DNA in the sample, e.g., via a 5' primer, allows the  
13 DNA to be attached to a solid support, leaving its 3' end remote from the support and  
14 available for subsequent hybridization with the extension primer and extension by the  
15 polymerase (or ligase). Alternatively, an unamplified DNA sample, such as a vector  
16 or a biological sample, may include, or be modified to include, a functional group that  
17 allows attachment to a solid support. In a related embodiment, the vector may include  
18 a means for attachment to a solid support adjacent to the site of insertion of the  
19 sample DNA such that the amplified DNA sample and the means for attachment may  
20 be excised together.

21 The solid support may conveniently take the form of, for example, microtiter  
22 wells, a solid support activated with polystyrene to bind the DNA sample (e.g., primer  
23 DNA), particles, beads (e.g., nylon beads, polystyrene microbeads, or glass beads)  
24 (Polysciences, Warrington, PA), glass surfaces, plates, dishes, flasks (Corning Glass

1 Works, Corning, NY), meshes (Becton Dickinson, Mountain View CA), membranes  
2 (Millipore Corp., Bedford, MA), dipsticks, capillaries, hollow fibers (Amicon  
3 Corporation, Danvers, MA), screens and solid fibers (Edelman et al., U.S. Patent No.  
4 3,843,324; see also Kuroda et. al., U.S. Patent No. 4,416,777, incorporated herein by  
5 reference), or needles, made, for example, of agarose, cellulose, alginate, Teflon, or  
6 polystyrene. Magnetic particles, e.g., magnetic beads, may also be used as solid  
7 supports, and such materials are commercially available (Robbin Scientific, Mountain  
8 View, CA).

9       The solid support may alternatively or additionally carry functional groups  
10 such as hydroxyl, carboxyl, aldehyde, or amino groups, or other moieties, such as  
11 avidin or streptavidin, for the attachment of the appropriately modified DNA, e.g., via  
12 modified oligonucleotide primers used in an amplification reaction. These may in  
13 general be provided by treating the support to provide a surface coating of a polymer  
14 carrying one of such functional groups, e.g., polyurethane together with a polyglycol  
15 to provide hydroxyl groups, or a cellulose derivative to provide hydroxyl groups, a  
16 polymer or copolymer of acrylic acid or methacrylic acid to provide carboxyl groups,  
17 or an aminoalkylated polymer to provide amino groups. Various other supports and  
18 methods of attachment and detachment of nucleic acid molecules to supports, with  
19 and without the use of a linker, is described in U.S. Patent No. 5,789,172,  
20 incorporated herein by reference.

21       As indicated above, the DNA sample need not be attached to a solid support.  
22 For example, a polymerase extension reaction may be carried out in solution on a  
23 DNA sample that is prepared in the context of a primer extension reaction having a  
24 buffer that will accommodate the addition of an oligonucleotide primer, a DNA

1 polymerase, cdNTPs, and a single or double-stranded DNA template. A ligation  
2 extension reaction may be similarly carried out in an appropriate buffer in the  
3 presence of an oligonucleotide primer, a DNA ligase, tagged oligonucleotide, and a  
4 single or double-stranded DNA template.

5

#### 6 *Extension*

7       Once a suitable DNA sample is prepared, the sample is subject to a primer  
8 extension reaction by addition of an oligonucleotide primer, a DNA polymerase, and  
9 four cdNTPs, such that one base is incorporated onto the DNA template before  
10 extension is blocked by the cleavable tag on the incorporated base. Alternatively, an  
11 oligonucleotide ligation reaction is used to extend the template DNA sample, as  
12 described above. Those skilled in the art will appreciate that such extension reactions  
13 can be modified to accommodate variations in template DNAs, reaction conditions,  
14 etc. It will be further recognized that the chosen oligonucleotide primer must be  
15 sufficiently large to provide appropriate hybridization with the target DNA sequence.  
16 Moreover, the oligonucleotide primer preferably hybridizes immediately 5' to the  
17 target sequence. Guidance for selection of primers and primer extension reactions can  
18 be found in the scientific literature, for example, Maniatis et al., *Molecular Cloning, a*  
19 *laboratory Manual* (1989).

20       The polymerase in the primer extension reaction may be any polymerase that  
21 incorporates dNTPs, and preferably cdNTPs, onto a single stranded DNA template.  
22 Examples of suitable polymerases that may conveniently be used, and many are  
23 known in the art and reported in the literature, include T7 polymerase, Klenow, and  
24 Sequenase. Those skilled in the art will be aware that certain polymerases, e.g., T7

1 polymerase, recognize a specific leader sequence in the DNA, which can be included  
2 in the sequence of the oligonucleotide primer. If a double stranded DNA template is  
3 to be used in the polymerase extension reaction, it is desirable that a thermostable  
4 polymerase, such as a Taq polymerase, be chosen to permit repeated temperature  
5 cycling without having to add additional polymerase for each round of extension.

6 It is well known that many polymerases have a proof-reading or error  
7 checking ability, which sometimes results in digestion of 3' ends available for  
8 extension. In the method of the invention, such digestion may result in an increased  
9 level of background noise. In order to avoid this problem, a nonproof-reading  
10 polymerase, e.g., an exonuclease deficient (exo-) Klenow polymerase may be used.  
11 Otherwise, fluoride ions or nucleotide monophosphates that suppress 3' digestion by  
12 the polymerase may be added to the extension reaction mixture. In addition, it may  
13 be advantageous to use an excess amount of polymerase over primer/template to  
14 maximize the number of free 3' ends that are extended. Those skilled in the art will  
15 appreciate that the precise reaction conditions and concentrations of reactants etc. may  
16 readily be determined for each system according to choice.

17 Since the primer is extended by a single base (or a single oligonucleotide) by  
18 the methods described above, the extended primer serves in exactly the same way in  
19 the repeated procedure, and with each subsequence base (or oligonucleotide) addition,  
20 to determine the next base or bases in the sequence, permitting the whole sample to be  
21 sequenced.

22

1     *Separation*

2             In the case of the polymerase extension reaction, prior to cleavage of the tag  
3     from the extended DNA template, the excess cdNTPs must be removed from the  
4     reaction mixture to prevent contamination of the cleavage product with signals from  
5     other unincorporated bases. As mentioned above, this separation may be  
6     accomplished by washing the cdNTPs through a membrane filter that allows flow  
7     through of small molecules such as water, salts, and cdNTPS, but does not allow the  
8     flow through of larger molecules such as the polymerase and the DNA template.

9             In the case of the ligase extension reaction, prior to cleavage of the tag from  
10    the oligonucleotide on the extended DNA template, the unincorporated tagged  
11    oligonucleotide must be removed from the reaction mixture to prevent contamination  
12    of the cleavage product with signals from the unincorporated tagged oligonucleotides.  
13    Depending on whether the DNA template is free in solution or attached to a solid  
14    support, the excess unincorporated tagged oligonucleotide may be removed by either  
15    filtration or washing the solid support, respectively. Those skilled in the art will  
16    appreciate that if the DNA ligase is removed from the extension reaction mixture  
17    along with the tagged oligonucleotide, the ligase will need to be added back to the  
18    extension reaction mixture in subsequent rounds. This, of course, is also applicable to  
19    a sequencing reaction that utilizes a polymerase, where the polymerase is removed  
20    from the extension reaction in a separate step with the cdNTPs.

21            Those skilled in the art will further appreciate that a wide variety of membrane  
22    filters are available in the art. For example, molecular filtration, also known as  
23    ultrafiltration, is a membrane separation technique used to segregate substances  
24    according to molecular weight and size. Molecular filtration is ideally suited to

1 separate salts and other low molecular weight solutes from high molecular weight  
2 species. Molecular filtration is based on a pressure differential across the  
3 semipermeable membrane to drive permeable materials through the membrane. For  
4 this reason, molecular filtration typically separates solutes and concentrates retained  
5 materials more rapidly. Molecular filtration membranes appropriate for use in the  
6 present invention may be purchased from Millipore Corp., Bedford, MA.

7 In another preferred embodiment, a flow through cell is used for single  
8 stranded DNA analysis. In this embodiment, the tag is washed away and is sent to the  
9 detector directly. One example of a variation on the flow through cell approach that  
10 would be amenable to multiplexing is to use a 96 well plate with an ultrafiltration  
11 membrane incorporated in the well. The excess reagents are either washed through by  
12 pressure or centrifuged through. The tag is then subsequently cleaved from extended  
13 nucleotide base, washed through the membrane, and collected for analysis by the  
14 method appropriate for the type of tag to be identified. In certain preferred  
15 embodiments, the different wells are pooled and the tags analyzed simultaneously to  
16 provide greater sample multiplexing as well as throughput.

17 Where the DNA template is immobilized on a solid support, the separation is  
18 accomplished by simply washing the cdNTPs (or tagged oligonucleotide) away from  
19 the solid support. For example, one basic approach to retaining the DNA for analysis  
20 would be to absorb the target DNA to an adsorptive surface instead of trapping it  
21 behind an ultrafiltration membrane. The excess reagents are washed away from the  
22 absorbed DNA by rinsing the adsorptive surface with a wash solution. The solvents  
23 used in the wash step must be chosen to avoid loss of the DNA during the wash steps.

1           The basic concept of using a membrane to permit flow of the excess reagent  
2 away from the DNA in the wash step can be further combined with the concept of  
3 adsorbing the DNA to a surface by incorporating a membrane onto a microfluidic  
4 chip. Solvent addition, or washes, may be carried out by the use of electro-osmotic  
5 flow. In this particular embodiment, all of the reactions and sample pooling occurs on  
6 the chip, permitting high throughput at a lower cost compared to the well plate  
7 approach. Within further embodiments, the steps of removing, cleaving, and  
8 detecting may be performed in a continuous manner (e.g., as a continuous flow), for  
9 example, on a single device which may be automated.

10

#### 11 *Cleavable Tags and Detection*

12           A "tag," according to the present invention, is a chemical moiety that is used  
13 to uniquely identify a nucleic acid molecule. In certain preferred embodiments, the  
14 nucleic acid molecule is a nucleotide base. In other preferred embodiments, the  
15 nucleic acid molecule is a nucleic acid fragment, such as a DNA or an RNA. "Tag"  
16 more specifically refers to the tag variable component as well as whatever may be  
17 bonded most closely to it.

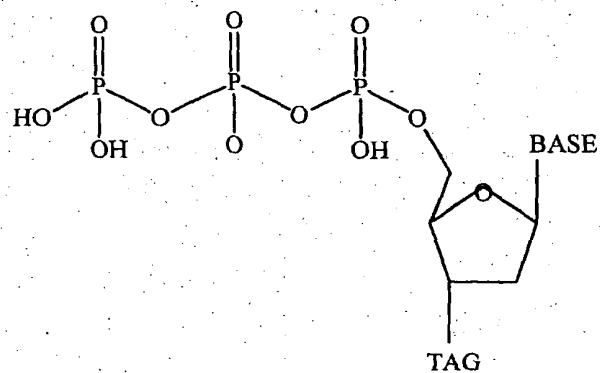
18           The tags of the present invention further possess one or more of certain  
19 characteristic attributes. The tag is preferably distinguishable from all other tags,  
20 particularly from other tags used in a particular reaction. The discrimination from  
21 other chemical moieties can be based on the chromatographic behavior of the tag  
22 (particularly after the cleavage reaction), its spectroscopic or potentiometric  
23 properties, or some combination thereof. In addition, the tag is capable of being  
24 detected when present at

1  $10^{-22}$  to  $10^{-6}$  mole. The tag is further attachable to the nucleic acid molecule, e.g.,  
2 nucleotide base or oligonucleotide, through a "chemical handle" (see U.S. Patent No.  
3 6,027,890, incorporated herein by reference) which may attach the tag to the nucleic  
4 acid molecule either directly, or through a linker group. In certain preferred  
5 embodiments, the tags block primer extension. The tags are further stable toward all  
6 manipulations to which they are subjected, including attachment to the nucleic acid  
7 molecule and cleavage from the nucleic acid molecule, and any manipulations of the  
8 nucleic acid molecule while the tag is attached to it; nor does the tag significantly  
9 interfere with the manipulations performed (e.g., hybridization or enzymatic  
10 reactions) on the nucleic acid molecule while the tag is attached to it.

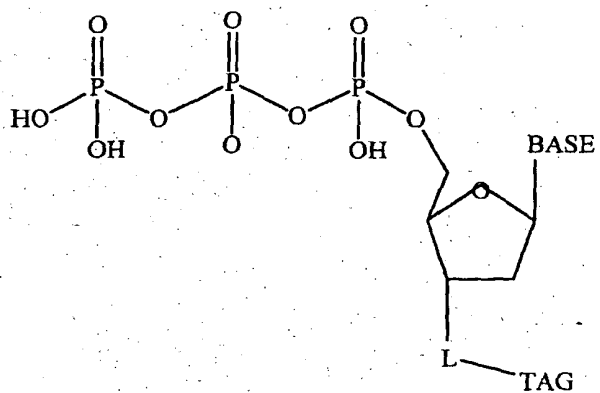
11 The tags of the present invention include any tag that is cleavable by chemical  
12 means or by light, and such tags are discussed in detail below. Chemically cleavable  
13 tags include tags that are cleavable by an acid or a base. Photo-cleavable tags include  
14 tags that are cleavable by a wavelength of light. Other methods of cleavage include  
15 oxidation, reduction, enzymatic, electrochemical, heat, and the like.

16 As mentioned above, the tag is further capable of terminating a primer  
17 extension reaction. In certain preferred embodiments, the terminating nature of the  
18 tag may be due to the nature of the tag itself, for example the structure of the tag, e.g.,  
19 a tag that is sufficiently bulky in its structure so that that it prevents addition of any  
20 additional bases to the extension product. Alternatively or additionally, the  
21 terminating nature of the tag may be due to the placement of the tag on the base.  
22 Preferably, the tag is attached to the base so that when the base is added to the  
23 growing 3' end of the extension product the tag effectively blocks the extension of the  
24 3' end by additional bases, once a tagged base has been added. One such example of

1 a tagged base, wherein the tag is attached directly to the base, that would block  
2 extension is shown below.



6 Alternatively, the tag is linked via a labile bond (or labile bonds) to the 3' position of  
7 the dNTP, as shown below,



11 wherein:

12 L is the linker.

1 According to the invention, the tag, including the linker in cases where a linker is  
2 employed, or other 3' blocking group, are removed to expose the 3' hydroxyl group of  
3 the base. Exemplary tags and linkers are described in detail in U.S. Patent 6,027,890,  
4 incorporated herein by reference.

5 In light of the availability of numerous tags, any number of tags may be  
6 utilized in a given reaction simultaneously, or within different reactions in an array.  
7 In certain preferred embodiments, particularly with respect to detection of ligation  
8 products, as described below, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90,  
9 100, 150, 200, 250, 300, 350, 400, 450, or greater than 500 different and unique  
10 tagged molecules may be utilized within a given reaction simultaneously, wherein  
11 each tag is unique for a selected base, oligonucleotide, or other nucleic acid fragment.

12 The characteristics of a variety of well known tags that are amenable to  
13 attachment to the bases and nucleic acid molecules of the invention are described in  
14 U.S. Patent 6,027,890, incorporated herein by reference. Such tags are detectable,  
15 once cleaved from the extended base, by fluorometry, mass spectrometry (MS);  
16 infrared (IR) spectrometry, ultraviolet (UV) spectrometry, or potentiostatic  
17 amperometry (e.g., utilizing coulometric or amperometric detectors). Mass  
18 spectrometry is particularly amendable to multiplexing with mass detection.  
19 Representative examples of suitable mass spectrometric techniques include time-of-  
20 flight mass spectrometry, quadrupole mass spectrometry, magnetic sector mass  
21 spectrometry, and electric sector mass spectrometry. Specific embodiments of such  
22 techniques include ion-trap mass spectrometry, electrospray ionization mass  
23 spectrometry, ion-spray mass spectrometry, liquid ionization mass spectrometry,  
24 atmospheric pressure ionization mass spectrometry, electron ionization mass

1 spectrometry, fast atom bombard ionization mass spectrometry, MALDI mass  
2 spectrometry, photoionization time-of-flight mass spectrometry, laser droplet mass  
3 spectrometry, MALDI-TOF mass spectrometry, APCI mass spectrometry, nano-spray  
4 mass specrometry, nebulised spray ionization mass spectrometry, chemical ionization  
5 mass spectrometry, resonance ionization mass spectrometry, secondary ionization  
6 mass spectrometry, and thermospray mass spectrometry.

7       The following is a list of representative vendors for separation and detection  
8 technologies that may be used in the present invention. Perkin Elmer/Applied  
9 Biosystems Division (ABI, Foster City, CA) manufacturers semi-automated  
10 sequencers based on fluorescent-dyes (ABI373) and (ABI377). Analytical Spectral  
11 Devices (Boulder, CO) manufactures UV spectrometers. Hitachi Instruments (Tokyo,  
12 Japan) manufactures Atomic Absorption spectrometers, Fluorescence spectrometers,  
13 LC and GC Mass Spectrometers, NMR spectrometers, and UV-VIS Spectrometers.  
14 Perseptive Biosystems (Framingham, MA) produces Mass Spectrometers (Voyager™  
15 Elite). Bruker Instruments Inc. (Manning Park, MA) manufactures FTIR  
16 Spectrometers (Vector 22), FT-Raman Spectrometers, Time of Flight Mass  
17 Spectrometers (Reflex II™), Ion Trap Mass Spectrometer (Esquire™) and a MALDI  
18 Mass Spectrometer. Analytical Technology Inc. (ATI, Boston, MA) makes UV  
19 detectors and Diode Array Detectors. Teledyne Electronic Technologies (Mountain  
20 View, CA) manufactures an Ion Trap Mass Spectrometer 3DQ Discovery™ and the  
21 3dQ Apogee™). Perkin Elmer/Applied Biosystems Division, (Foster City, CA)  
22 manufactures a Sciex Mass Spectrometer (triple quadrupole LC/MS/MS, the API  
23 100/300), which is compatible with electrospray. Hewlett-Packard (Santa Clara, CA)  
24 produces Mass Selective Detectors (HP 5972A), MALDI-TOF Mass Spectrometers

1 (HP G2025A), Diode Array Detectors, CE units, HPLC units (HP1090), as well as  
2 UV Spectrometers. Finnigan Corporation (San Jose, CA) manufactures mass  
3 Spectrometers (magnetic sector and four other related mass spectrometers). Rainin  
4 (Emeryville, CA) manufactures HPLC instruments.

5 Those skilled in the art will recognize how to apply such devices to the  
6 methods of the present invention. Those skilled in the art will further appreciate that  
7 devices used to detect pyrosequencing reactions may be adapted to detect and identify  
8 the cleaved tags of the invention. For example, the reaction monitoring system  
9 described in WO 99/66131, the microfluidic device described in WO 00/40750, the  
10 liquid dispensing apparatus described in WO 00/56455, the solid support apparatus of  
11 U.S. Patent No. 5,302,509, each of which is incorporated herein by reference, may be  
12 adopted for use with the method of the present invention.

13  
14 *Automation and High-throughput Sequencing*

15 The DNA sequencing methods of the present invention are fully automatable.  
16 Those skilled in the art will recognize that the use of a robot apparatus, where a large  
17 number of samples may be rapidly analyzed, may be used for rapid detection and  
18 quantification of the tag molecules. Tags to be detected spectrophotometrically may  
19 be detected, e.g., by mass spectrometry or fluorescence spectrometry. The use of  
20 luminometers, mass spectrometers, and other spectrophotometric devices are well  
21 known in the art and described in the literature. The DNA sequencing method of the  
22 present invention thus provides an automated approach for high-throughput, non-  
23 electrophoretic sequencing procedures that allows for continuous measurement of the  
24 progress of the polymerization reaction in real time.

1           In related embodiments, it will be appreciated that multiple samples may be  
2   handled in parallel and such parallel handling provides another advantage to the  
3   inventive method. In order to obtain high throughput sequence readout, multiple  
4   DNA sequencing reactions can be processed in parallel. According to this particular  
5   embodiment, the DNA sequencing method of the present invention can be carried out  
6   in any of a variety of array formats.

7           For example, a single sequencing reaction of the invention, carried out in a  
8   single well and analyzed using flow injection analysis (FIA) has a rate of about one  
9   base every six seconds (equivalent to about ten bases per minute and about 600 bases  
10   per hour). In order to increase this rate, the DNA sequencing reactions may be  
11   multiplexed. For example, multiplexing 25 sequences increases the rate of  
12   sequencing to about 15000 bases per hour. Those skilled in the art will recognize the  
13   power of multiplexing as it is applicable to any means of detection described herein.  
14   The number of DNA samples that can be multiplexed for parallel analysis can range  
15   10 to 100, in some cases 100-500, and in yet some other cases, 100-1000 or more  
16   DNA samples.

17           In certain preferred embodiments, an array format is used for analysis wherein  
18   the DNA samples are distributed over a surface, for example, a microfabricated chip,  
19   thereby immobilizing an ordered set of samples in a 2-dimensional format. This  
20   allows the analysis of many samples in parallel. According to this embodiment of the  
21   invention, the DNA samples are arrayed onto any of a variety available microchips  
22   prior to commencing the sequencing reaction. Methods of producing and analyzing  
23   DNA arrays are well known in the art and are provided in U.S. Patent No. 6,027,789,  
24   incorporated herein by reference.

1           For example, applying the method of the invention to the array format, after  
2   primer extension, the tags may be cleaved from the DNA samples on the chip and  
3   pooled for analysis using spectrometric or potentiometric techniques (e.g., MALDI-  
4   MS). In one particular embodiment of the present invention, an array interrogation  
5   system is provided that includes a DNA array generating device, a washing device, a  
6   tag cleaving device, a detecting device, and a data processor and analyzer that  
7   analyzes data from the detecting device to correlate a tag with a nucleic acid fragment  
8   from a sample, as described in U.S. Patent No. 6,027,789, incorporated herein by  
9   reference. The arrayed DNA chip has on its surface selected DNA samples of nucleic  
10   acid fragments and cleavable tags, e.g., cleavable mass spectrometer tags, attached to  
11   the nucleic acid fragments. The arrayed DNA chip is passed through or past a  
12   photolytic cleavage device that cleaves the tags from the nucleic acid fragments while  
13   still on the DNA chip.

14           After the tags are cleaved, the DNA chip is positioned in an automated micro-  
15   array sampling laser device, such as a Matrix Assisted Laser Desorption Ionization  
16   (MALDI) instrument. The MALDI instrument is adapted to irradiate and cause  
17   desorption of the tags, which are transferred to a detection device, such as a mass  
18   spectrometer, wherein tags are identified based upon the difference in molecular  
19   weight.

20           Data from the detection device is provided to the data processor and analyzer,  
21   which includes a software program that maps the signature of a given tag to a specific  
22   sample. The software is able to display the DNA sequence determined and load the  
23   sequence information into respective data bases.

1 In an alternative embodiment, the MALDI instrument includes an additional  
2 light source that is capable of irradiating the entire DNA chip at a wavelength in the  
3 range of 250-360 nm with adjustable intensity, so as to cause the photolytic cleaving  
4 of the tags. Accordingly, the cleaving device is incorporated as a component of the  
5 MALDI instrument. After cleaving the tags, the MALDI instrument volatilized the  
6 tags, which are transferred to the detecting device as discussed above.

7 In yet another embodiment, the DNA chip is moved from the DNA array  
8 generating device directly to the MALDI instrument. The MALDI instrument  
9 includes a laser that emits at a wavelength in the range of approximately 250 to 360  
10 nm, inclusive. The laser causes the simultaneous photolytic cleavage of the tag from  
11 the nucleic acid fragment along with simultaneous desorption of the tag. The tags are  
12 then transferred to the mass spectrometer or other detection device, as discussed  
13 above. Accordingly, this alternate embodiment provides photocleavage by the  
14 MALDI instrument, so that a separate cleavage device is not needed.

15 If fluorescence sensing is employed in the present invention for detection of  
16 the tag, this increases the rate of the sequencing to one base every fifteen seconds  
17 (equivalent to about four bases per minute). If 100 sequencing reactions are arranged  
18 onto 100 lanes of the chip this yields a rate of about 24000 bases per hour. Similar  
19 sequencing rates are achievable with varying cleavage means.

20 Florescent tags can be identified and quantitated most directly by their  
21 absorption and fluorescence emission wavelengths and intensities. While a  
22 conventional spectrofluorometer is extremely flexible, providing continuous ranges of  
23 excitation and emission wavelengths ( $I_{EX}$ ,  $I_{S1}$ ,  $I_{S2}$ ), more specialized instruments, such  
24 as flow cytometers and laser-scanning microscopes require probes that are excitable at

1 a single fixed wavelength. In contemporary instruments, this is usually the 488-nm  
2 line or the argon laser.

3 Radioactive tags may also be applicable to the present invention. Radioactive  
4 tags may be detected by, e.g., a CCD detector.

5 In using fluorescent and radioactive tags, the number of different reactions that  
6 are simultaneously detectable may be more limited than, e.g., mass tags. For  
7 example, the use of four fluorescent molecules, such as commonly employed in DNA  
8 sequence analysis, limits analysis to four samples at a time.

9 In certain preferred embodiments, the sample reactions may be pooled on at  
10 least one array and the products detected simultaneously. By using a cleavable tag,  
11 such as the ones described herein, having a different molecular weight or other  
12 physical attribute in each reaction, the entire set of reaction products can be harvested  
13 together and analyzed.

14

#### 15 *Applications*

16 The invention in the above embodiments provides a simple and rapid method  
17 for sequencing a DNA sample. The methods of the invention both avoid the  
18 requirement of separation of the extension product and allows rapid, real-time  
19 analysis of the extension reaction. These methods have many applications, which will  
20 readily be appreciated by the skilled artisan.

21 To name but a few, the present invention is applicable in the field of forensics  
22 (e.g., the identification of individuals and the level of DNA sequence variations);  
23 tumor diagnosis (e.g., for detection of viral or cellular oncogenes in a biological  
24 sample from a patient); transplantation analyses (e.g., the identification of antigen

1 specific variable DNA sequences from a biological sample); diagnosis of autoimmune  
2 diseases, such as juvenile diabetes, arteriosclerosis, multiple sclerosis, rheumatoid  
3 arthritis, and encephalomyelitis; genome diagnostics (e.g., the identification of genetic  
4 defects or hereditary and acquired genetic diseases in newborns and adults, for  
5 example, schizophrenia, manic depression, epilepsy, sickle-cell anemia, thalassemias,  
6  $\alpha$ 1-antitrypsin deficiency, Lesch-Nyhan syndrome, cystic fibrosis, Duchenn/Becker  
7 muscular deficiency, Alzheimer's disease, X-chromosome-dependent mental  
8 deficiency, and Huntingtins chorea); infectious disease (e.g. detection of viral or  
9 microbial infection of a biological sample); mutation detection (e.g., detection of a  
10 mutated base in a DNA sample from a biological or artificial source); detection of  
11 single nucleotide changes (e.g., a primer hybridizes to a sequence adjacent to a known  
12 single nucleotide polymorphism and a cdNTP added to the adjacent position is  
13 detected and identified).

14 As mentioned above, the method of the present invention may be adapted for  
15 use with a ligase instead of a polymerase. One adaptation of this technique is to the  
16 oligonucleotide ligation assay, which is used to identify known sequences in very  
17 large and complex genomes. To elaborate briefly on the ligase extension reactions  
18 described above, the basis of this assay is the ability of a ligase to covalently join two  
19 diagnostic oligonucleotides as they hybridize adjacent to one another on a given DNA  
20 target. If the sequences at the probe junctions are not perfectly base-paired, the  
21 probes will not be joined by the ligase. When tags are used, they are attached to the  
22 oligonucleotide, which is ligated to the DNA sample. After a ligation is complete, the  
23 tag is cleaved and detected by any of the means described herein (e.g., mass

1 specrometry, infrared spectrophotometry, potentiostatic amperometry, or UV/visible  
2 spectrophotometry).

3 In certain preferred embodiment, the DNA sample is amplified prior to  
4 exposure to the oligonucleotide ligation assay.

5

#### 6 *Kits*

7 The present invention further provides kits for use in methods of the invention  
8 that contain at least the following reagents: a) an oligonucleotide primer suitable for  
9 primer extension of a particular DNA template; b) four cdNTPs of adenine, guanine,  
10 thymine, and cytosine bases; c) a polymerase; d) a separation means to separate  
11 unincorporated dNTPS from the extended DNA template; and e) a cleavage means.

12 In certain embodiments of the invention, a detection means will be provided.

13 However, the detection means may often be provided by the purchaser.

14 In alternative embodiments, if the kit is used for a ligation sequencing reaction  
15 assay it may contain at least a) an oligonucleotide primer suitable for primer extension  
16 of a particular DNA template; b) at least one tagged oligonucleotide; c) a ligase; d) a  
17 separation means to separate unincorporated oligonucleotides from the extended DNA  
18 template; and e) a cleavage means. The kit may further provide a detection means.

19 However, the detection means may also be provided by the purchaser.

20 Other embodiments of the invention will be apparent to those skilled in the art  
21 from a consideration of the specification or practice of the invention disclosed herein.

22 It is intended that the specification be considered as exemplary only, with the true  
23 scope and spirit of the invention being indicated by the following claims.